



Globus Toolkit

Enhancing and Supporting GridFTP: An Essential Component of DOE High-speed Networking

Steve Tuecke

Deputy Director, Computation Institute

Argonne National Laboratory and University of Chicago



Project Overview

- **Project title:**
Enhancing and Supporting GridFTP:
An Essential Component of DOE High-speed Networking
- **PI:** Steve Tuecke
- **Project start:** October 2011
- **Duration:** 3 years
- **Objectives:**
 - Enhancements to GridFTP protocol & Globus GridFTP
 - Support Globus GridFTP for DOE community



What is GridFTP?

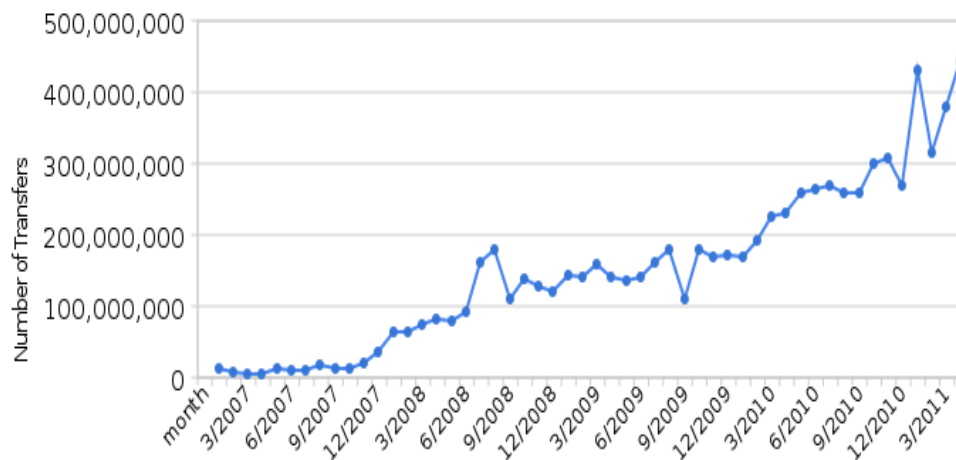
- **Standard workhorse for large data movement in distributed science projects across DOE and worldwide**
- **GridFTP protocol extends FTP for:**
 - High-performance
 - Strong security
 - Reliability
- **Globus GridFTP server**
 - Mature, widely used implementation of GridFTP



Globus GridFTP Usage

- **>5,000 Globus GridFTP servers installed**
- **>1/2 PB per day transferred**
- **Used by ESnet, NERSC, LCFs, APS, OSG, Globus Online, XSEDE, EGI, ...**

Monthly Totals* of GridFTP File Transfers



*for those "reporting"

Monthly Totals* of TBs Transferred Via GridFTP



*for those "reporting"



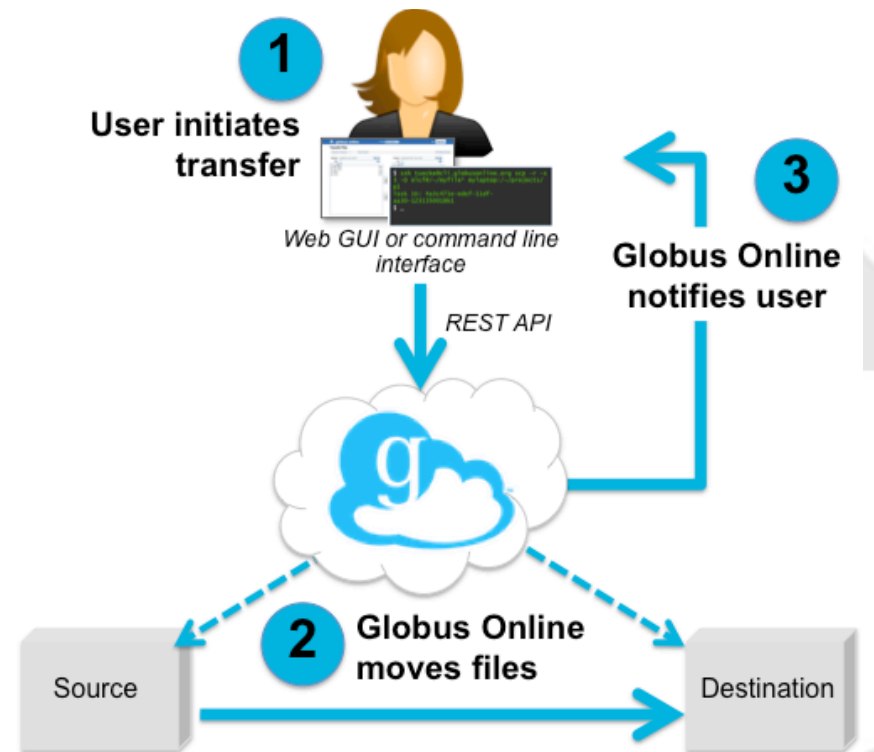
DOE use cases driving GridFTP

- **Globus Online driving GridFTP transfers**
 - NERSC: Recommended method for transferring files to/from GPFS file systems and HPSS
 - ALCF: User remote file transfer to/from GPFS
 - Advanced Photon Source: Distributed instrument data to users worldwide
 - ESnet DTNs
 - Etc.
- **Science communities with custom clients**
 - E.g. HEP, ESG



What is Globus Online?

- **Reliable file transfer.**
 - Easy “fire-and-forget” transfers
 - Automatic fault recovery
 - High performance
 - Across multiple security domains
- **No IT required.**
 - Software as a Service (SaaS)
 - No client software installation
 - New features automatically available
 - Consolidated support & troubleshooting
 - Works with existing GridFTP servers
 - Globus Connect solves “last mile problem”
- **>3500 registered users, >3PB moved**
- **Recommended by NERSC, ALCF, XSEDE, Blue Waters, and many campuses**





Enhancing and Supporting GridFTP Project Objectives

- **Leverage next-generation Terabit networks and multi-core processors**
 - Support big data transfer needs of exascale computers and scientific instruments
 - Not just big files, but lots of small files (LOSF), and end-to-end checksum verification
- **Improve support for firewalls and NATs**
- **Simplify ease of use and administration**
- **Support DOE facilities and scientists**



Proposed Work

- **Performance improvements**
 - Parallel command execution (checksums and LOSF)
 - Enable multiplexed transfers
 - Efficient recursive directory transfers
- **New firewall friendly data channel mode(s)**
 - Fix mode E connection directionality limitation
 - Single port GridFTP; no ephemeral ports
 - TCP simultaneous open, UDP NAT traversal
- **Simplify installation and configuration**
- **Native Windows implementation**
- **Better UDT support**



Supporting DOE

- **Fixed HPSS DSI plugin for NERSC**
- **Added server configuration to restrict access to specific paths (ALCF)**
- **Added data channel security for SSH GridFTP**
- **Fixed special character handling in local group names (APS)**



Simplified install & config

- **GT 5.2 native Linux packaging (RPM, dpkg)**
- **Native Windows implementation (single-user)**
- **Globus Connect Multi-User (GCMU)**
 - Simple install of GridFTP for use with Globus Online



Firewall & NAT friendliness

- **Added MLSC directory listing over control channel (no data channel or ephemera ports)**
- **Drafting requirements for protocol changes**
 - Single port GridFTP; no ephemeral ports
 - Both endpoints behinds NATs
 - Heartbeat all long commands to avoid timeouts
- **Investigating potential solutions:**
 - TCP-based data channel over single port
 - TCP simultaneous connect
 - UDP NAT traversal with STUN & ICE



Improved end-to-end performance

- **Drafting requirements for protocol changes**
 - Lots of small files
 - Parallel (out-of-order) execution
 - Multiplex files on data channel (help mask I/O latency)
 - Efficient recursive directory transfers
 - Efficient checksums of files
 - Parallel (out-of-order) execution
 - Algorithms
 - Mass storage system optimizations
 - Staging control
 - File ordering hints